



OIPE

RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/09/804,409B

TIME: 16:19:06

Input Set : A:\0278721.APP.txt

Output Set: N:\CRF4\06252003\I804409B.raw

RECEIVED
SEP 22 2003
TECH CENTER 1600/2900

3 <110> APPLICANT: ENGINE, INC.
4 <120> KIEFFER, TIMOTHY J.
5 <130> CHEUNG, ANTHONY T.
6 <140> TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATED PROTEIN
7 <150> EXPRESSION IN GUT
8 <160> FILE REFERENCE: 029996/027 8721
9 <170> CURRENT APPLICATION NUMBER: 09/804,409B
10 <180> CURRENT FILING DATE: 2001-03-12
11 <190> NUMBER OF SEQ ID NOS: 19
12 <210> SOFTWARE: PatentIn Ver. 2.1
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 19
15 <212> TYPE: DNA
16 <213> ORGANISM: Artificial Sequence
17 <220> FEATURE:
18 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
19 <400> SEQUENCE: 1
20 ccagccgcag ccttctga 19
21 <210> SEQ ID NO: 2
22 <211> LENGTH: 21
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
27 <400> SEQUENCE: 2
28 ggtccagcat tgttcacaa tg 22
29 <210> SEQ ID NO: 3
30 <211> LENGTH: 19
31 <212> TYPE: DNA
32 <213> ORGANISM: Artificial Sequence
33 <220> FEATURE:
34 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
35 <400> SEQUENCE: 3
36 accaccagcc ctactgat 19
37 <210> SEQ ID NO: 4
38 <211> LENGTH: 21
39 <212> TYPE: DNA
40 <213> ORGANISM: Artificial Sequence
41 <220> FEATURE:
42 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
43 <400> SEQUENCE: 4
44 ctactgatgac tactctcca gc 22
45 <210> SEQ ID NO: 5

RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/09/804,409B

TIME: 16:19:06

Input Set : A:\0278721.APP.txt

Output Set: N:\CRF4\06252003\I804409B.raw

68 <211> LENGTH: 1319

69 <212> TYPE: DNA

70 <213> ORGANISM: Mus musculus

71 <400> SEQUENCE: 5

```

73 ccgaaattac ccactaagtt ggaattctat aagggttggg tttgctgttt tgtttacagc 60
74 tccgtctttg gcacccagca cagctgagtg gttctaagcc caagtccatg cttaacacat 120
75 gtttgtttga tgaatacacg cgaagccggt tctcatttag gggcatgagt aggcagaggt 180
76 gtgggcagga agcaggaaag agcggaaaca ggtgoggaca gaaaggaggg gctctgaagg 240
77 atgcacgtca gtgcacaaat gtcacccaga taccaggttc actgtggccc taggccaggc 300
78 tgcacggggc tcccatgtg gtctgcccag ggtgagagca gaactggggt gggcggggca 360
79 gaagaaacc aaccaggaa gagggttgca cccaaattat ccaggtttta agtacattta 420
80 agagacaagg ctgggtgttt gaaggtcaga ggtgtccctg ggggtgtgga ctaggactga 480
81 ccactctgtt tttagtttaa tggtagaacc tgcctcacac tgctaactgc cttacttgcc 540
82 ccttgagagt tttagacctt ggacccccc atgtgtgggt tggaccttca gtcacacact 600
83 gaacgtgtgt gaagccactg gttgtccag aggggtctct ggcactgagg aagcagtga 660
84 cactatccc talcaataa caattaaata cacacagaat ggcaggcaca caactgagtt 720
85 tcaggagagg cctcgctcag gcaaggggtt caagaggctt ctgtgggacc cgttgatgt 780
86 tccaggaggt tcctaaagat ggggtgtgct ccagccaagt gaaatcaaga gaaaagtacg 840
87 cgaagtatag gaaactcag cagtctggag aggtaaatag gggaggaatc ctaggctcag 900
88 agacaggagt gaattgccc cggacgcaca gcaagttggc aggtggagtt cagctgtgoc 960
89 acctctgaa ggggggtacc ctttacagcc accagataca agcgggatag agacagctga 1020
90 tggagaagct ggaggtggg ggcgggaccc cgaaggtggg gaaagggggc ggggggggg 1080
91 tccatgaag taatttctg ggtgtgtgog cgtgtgtgog tgcgtgtgog tgtatataaa 1140
92 agcgggata gaattgtgog tgcgtgoccc gccacggcca ccatcacccc tgttaccacc 1200
93 accgtaact cagtgttccc gctgggtcag agctttggta gccagactac agaccactc 1260
94 ccgcatact cctgcagcag ctgctccact ctttcgcac cgtccggctc gctatgocg 1319

```

97 <210> SEQ ID NO: 6

98 <211> LENGTH: 1760

99 <212> TYPE: DNA

100 <213> ORGANISM: Mus musculus

101 <400> SEQUENCE: 6

```

103 gggaaatttc totagctctt tcattagggg cctgtgttcc catctaatag ctgactgtga 60
104 gcatacaatt ctgtgcttgc caggcactgg catagctcca caagagacag ctatatcagg 120
105 gtcttgtoag caaaatcttt ctggcatatg caatagtgtc tgggttttgt ggttgtatat 180
106 gggctggatc cccgggtggg gcagtctctg gatggctctt ccttcogtct tagctccaaa 240
107 ctttgtctct gtaactcctt ccattgggtac ttgtttccc attctaagaa ggagcaaaat 300
108 atccacaatt ccttctctct ccttctctct gagttttgca aatgcacaaa aactttcaaa 360
109 ccttctgaa tagccttctc tttagtgctt tccaatgtat attaaaataa tctatctttc 420
110 atccccattg attaaagcct tcttaaagcc agaaaactat attcattttt tctttttccc 480
111 agtagttcac aaactatctg gcacctcata agcatcataa ctacgttggg gggtagataa 540
112 aattgggaat tgattgttca gtcagcagag acttttagag gacctcatal aacaagattc 600
113 tctcagttct cagaaatata ttccagtata tacagggtta gaggactcac atctttaata 660
114 caataaaagt aaaaatttag acctgtataa attattaagg taactaatca agttccacgg 720
115 caaagtacag ccattgttat gaattataaa tccaagaagc ggtgggttaa ctctgacatt 780
116 gttccttggg tggttctcat tcattgaagt tagtcacctc aaacttactc accaaaacct 840
117 agaagtatct ctgtggtaac atgttctctt gatgcacaag gggctctagg catatgaaaa 900
118 tctctcaate tctctccctc tctctccccc ttccacccc actctctctc ttctagcagt 960
119 aatccctccc tctctggtag gcagtatgtt ttttgagaca cagtttctta gctatctctt 1020
120 gcaacacctg attttgtga agatttgaat ggctccatat agaagtatca acaacttgag 1080

```

RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/09/804,409B

TIME: 16:19:06

Input Set : A:\0278721.APP.txt

Output Set: N:\CRF4\06252003\I804409B.raw

```

121 cgtctgtgaa ctctcatttt gacactgtgc tgaaagaatt ggagttgatt ctcaataaaa 1140
122 aaaaaattaa gcatctcacc ttttttgctc aaactaaaac gttttaaaac agttctgcoct 1200
123 ggagtcatga tatgaaatac gatctatcat atttgcaatg ttctgttcaa ttgtggctgc 1260
124 accaggaaat gagaagctat ttctttatag gcacaaataa aaagatagtc attatctgta 1320
125 aaattcttat gacatggcag caagcccaag aaacctttct aaacaaggcg tgaaaacgca 1380
126 gagatgtcct tgcgaattag catgtctatc tgacagattt ctccctttct aagggaattt 1440
127 gtgctgaaca ttttatttct agcctcagag ataaaagaag ggggaagaag ctgtagtttt 1500
128 tgctacataa gacagggtgc gtaagcatgc aacgctttta aaaaatatct aaagtgattg 1560
129 ttttctctcg gattctttga aaaagctcgc ctgogctggg gtttgagggt gagccggtga 1620
130 cgtcagcgtg gaatggggag tcaggcgccc aggcctctct taagccgagg agctgtcccg 1680
131 tgotgaaaag gcccgagccc tcactcagcg gcagagagga gcctgcttgg agccttccac 1740
132 ataataaag acagaggtaa                                     1760
135 <110> SEQ ID NO: 7
136 <111> LENGTH: 2623
137 <112> TYPE: DNA
138 <113> ORGANISM: Mus musculus
140 <400> SEQUENCE: 7
141 agcttttaggt gtgtgaatat ctactttggt gctagggcoct tggtcatact aagtaagttt 60
142 ccccttcaact ggggtgtacc agtttacct ggactgtcta agcaacaaga aggatagaca 120
143 tggcctaaca cagatttcat gtctgccact ggctatgtca gaacatgtag gagcttttgg 180
144 aatcagtgaa acagggtatt tcagactgcc ttccctgggt ggggctttcc cgaagccata 240
145 ttttctctag agtcagcctt tcccagctga ggacaagctg taactggacag atgccaagcca 300
146 ctgtgaactgg gaatacatgg tcatttaggc agctggctta tctcatccat ggtacttgat 360
147 ggtctgggtt cagcacctca cagaaagtcc agacgggagg ctcccgagaa aacagagaag 420
148 caggcaggag atcctgcagg caatcctcct gctccacagc ctgcatggac ttccctcagc 480
149 cttagtgctg gtgggtccca ttgagaaca ttggttatat gttattttca aacgatctg 540
150 cctttaagga ttggaagaaa aaaactgtgg tgtttgggtt acctttatga taatggcctt 600
151 ttcatctctc taataaatat tgcgaagtag ggtagattct ataagaaaac tcttaaccca 660
152 tggatttagc aaatcatgta ggtgctaata atgaatactg gatgcagtca gtacagggat 720
153 ataaaatgga atgtaagagc ctgttgctat gaatggttag cttaactagat gttgtacaag 780
154 aaatgttgac gttatgact gtggaaaact ggtattgaag atgtggactc gaaaactttg 840
155 ggattttttg atgcatgat aaaaatgtga agaatactgt tcttaacca aagaagaag 900
156 aagaaggaga aggaggagga agaggaggag gaggaagaag agggggagga agaagaagag 960
157 aaggaggagg aaggaggagg ggaggagaa gaggaggagg aggaagaaga agagaaggag 1020
158 gaggaactag aggaggagga gaagaaggag aaggggaaag agagagtac cagaacattt 1080
159 ggggtgccc cagaatacca gatactccag acatagtcac agaaggactg gtttgtttgt 1140
160 taaataggtg ctttgaaaag ttgtggggga aacctgcagt gagatttgtt gtcttagaaa 1200
161 tgataggcaa gattcatcca caagaatgc acaagatggc tgctgaaca agcctgaac 1260
162 attaacagca ccagtagacc tgcttacaag gaagaaagca atctcatagg cctccacccc 1320
163 aaacaaagac tacagacagc agaggaaact gagagcagga gaaattgggt ctccctttta 1380
164 tgagccccc aactggttgt caaatactca atggtcagcc ctgaaatcat atgcacaaa 1440
165 taatactagc gcaactgaa agattgtagc tgtgtgtgtg tgtgtaatga taacaaagaa 1500
166 gaaaaggccc catgttagag agggagcaag gtgggcattg aggtatggaa ggagtggaa 1560
167 ggaggggtga gaaggggaaa gtgatgtaat tatcttttaa ttataaaaa aataaaaaat 1620
168 gggctggtga gatggctcag tgggtaagag caccgcactg cttcttcoga aggtctggag 1680
169 ttcaaatccc agcaaccaca tgggtggtca caaccatcgg taacgagatc tggcgccctc 1740
170 tctggagtg totgaagaca gotacagtgt acttacatat aataaataaa taaatctttt 1800
171 aaaaaaata aaaaataaaa tattagaata aaatgtagag gaatatttt aatttaacaa 1860
172 ctgggtgtg gcaaaagctt tcttaacaa aaacttaac ctcagataa gaaaagacta 1920

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,409B

DATE: 06/25/2003

TIME: 16:19:07

Input Set : A:\0278721.APP.txt

Output Set: N:\CRF4\06252003\I804409B.raw

```

173 gaatccacga cgtggataga tactttctgt tgatgcaaga cactatttat caggttgtaa 1080
174 cttgagcaga acttgagttg taacttggtg ggaaacacaa cacccttggc aaacaaaga 2040
175 ttactagata ttttagatga aatataaaaa tactttccac aactgatagg taggaacag 3100
176 ttcaatagta atataattat tgaacaaata atccttaaaa gaagaaatcc agagggaatag 4160
177 caagttaggg gaagagaggg tgtgtgtgtg tgtgtgtgtg cgcacattta tagccaaaat 5220
178 agatgatata ctttaattgaa catgccaatta aaaccattta ttttgcatac agtttacata 6280
179 tgcataatgaa tacttaaaaa aaaaacattg ggatttgaga gaaatggctc agtggttaag 7340
180 atttcaatto ccagcaacca catgattgct cacaaccatc tgtaatggga totgatgoot 8400
181 tctttctgta tctctgaaga aagtgaacgt gtacttataa ttataaataa ataaatcttt 9460
182 aacccaaaaa ccccccataa ttcaacaaac gatatgtcct ggtctgagga ttccagggcat 10520
183 agaaatagaa acacacagag tgtggagaca gtgoggttca ggtccggcat tccagttcag 11580
184 gcttcagacc aagagaaagg gaaaaganga gacaagcaac aag 12640

```

187 <110> SEQ ID NO: 3

188 <111> LENGTH: 226

189 <112> TYPE: DNA

190 <113> ORGANISM: Homo sapiens

191 <400> SEQUENCE: 3

```

193 tccaggaaat ccggatccca ggcggggggg cggggggggg gctcgggga gagggggggc 60
194 ccccggaacg ccgggggggg gggggggggg cggggggggg ccggttaaga agagcgtggc 120
195 cggggggggg ccccgctggc cccagggaaa gcgagggggg ccccgagcgg gcagagaccc 180
196 ccccgagcgg cgggggggga ggcagggggg ggcgagcggg ggcacc 240

```

199 <110> SEQ ID NO: 4

200 <111> LENGTH: 110

201 <112> TYPE: PRT

202 <113> ORGANISM: Homo sapiens

203 <400> SEQUENCE: 4

```

205 Met Ala Leu Trp Met Arg Ser Leu Pro Leu Leu Ala Leu Leu Ala Leu
206      1              5              10              15
208 Trp Gly Pro Asp Pro Ala Ala Ala Phe Val Asn Gln His Leu Cys Gly
209      20              25              30
211 Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe
212      35              40              45
214 Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Leu Gln Val Gly
215      50              55              60
217 Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu
218      65              70              75              80
220 Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys Cys
221      85              90              95
223 Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
224      100              105              110

```

225 <110> SEQ ID NO: 10

226 <111> LENGTH: 450

227 <112> TYPE: DNA

228 <113> ORGANISM: Homo sapiens

229 <400> SEQUENCE: 10

```

233 gctccatcag aagaggccat caagcacatc actgtccttc tgcacaggcc ctgtggatgc 60
234 gctcctggcc cctgtgtggg ctgtgtggcc totggggacc tgaaccagcc gcagcctttg 120
235 tgaaccaaca cctgtgcggc tcaacagg gggaagctct ctacctagtg tggggggaac 180
236 gagccttctt ctacacaccc aagaccggcc gggaggcaga ggacctgcag gtggggcagg 240

```

RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/09/804,409B

TIME: 16:19:07

Input Set : A:\0278721.APP.txt

Output Set: N:\CRF4\06252003\I804409B.raw

```

237 tggagctggg cgggggccc ggtgcaggca gctgcagcc cttggccctg gaggggtccc 300
238 tgcagaagcg tggcattgtg gaacaatgct gtaccagcat ctgctccctc taccagctgg 360
239 agaactactg caactagaag cagcccgag gcagccccc acccgccgcc tctgcacgg 420
240 agagagatgg aataaagccc ttgaaccagc                                450

```

243 <210> SEQ ID NO: 11

244 <211> LENGTH: 167

245 <212> TYPE: PRT

246 <213> ORGANISM: Homo sapiens

248 <400> SEQUENCE: 11

```

249 Met His Trp Gly Thr Leu Cys Gly Phe Leu Trp Leu Trp Pro Tyr Leu
250   1             5             10             15
251 Phe Tyr Val Gln Ala Val Pro Ile Gln Lys Val Gln Asp Asp Thr Lys
252             20             25             30
253 Thr Leu Ile Lys Thr Ile Val Thr Arg Ile Asn Asp Ile Ser His Thr
254             35             40             45
255 Gln Ser Val Ser Ser Lys Gln Lys Val Thr Gly Leu Asp Phe Ile Pro
256             50             55             60
257 Gly Leu His Pro Ile Leu Thr Leu Ser Lys Met Asp Gln Thr Leu Ala
258             65             70             75             80
259 Val Tyr Gln Gln Ile Leu Thr Ser Met Pro Ser Arg Asn Val Ile Gln
260             85             90             95
261 Ile Ser Asn Asp Leu Glu Asn Leu Arg Asp Leu Leu His Val Leu Ala
262             100            105            110
263 Phe Ser Lys Ser Cys His Leu Pro Trp Ala Ser Gly Leu Glu Thr Leu
264             115            120            125
265 Asp Ser Leu Gly Gly Val Leu Glu Ala Ser Gly Tyr Ser Thr Glu Val
266             130            135            140
267 Val Ala Leu Ser Arg Leu Gln Gly Ser Leu Gln Asp Met Leu Trp Gln
268             145            150            155            160
269 Leu Asp Leu Ser Pro Gly Cys

```

280 165

283 <210> SEQ ID NO: 12

284 <211> LENGTH: 3403

285 <212> TYPE: DNA

286 <213> ORGANISM: Homo sapiens

288 <400> SEQUENCE: 12

```

289 tctgttttca ggcccagaag gcccattctg ggaaggaaaa tgcattgggg aacctgtgac 60
290 ggatttctgt ggctttggcc ctatcttttc tatgtccaag ctgtgcccat ccaaaaagtc 120
291 caagatgaca ccaaaacctt catcaagaca attgtcacca ggatcaatga catttcacac 180
292 aagcagtcag tctctcccaa acagaaaagtc aacgggtttg acttcattcc tgggtctcac 240
293 cccattctga ccttatccaa gatggaccag acaactggcag tctaccaaca gatctccacc 300
294 agtatgcctt ccagaaaagt gatccaaata tccaaacgacc tggagaaact ccgggatctt 360
295 cttcacgtgc tggccttctc taagagctgc cacttgcctt gggccagtgg cctggagacc 420
296 ttggacagcc tggggggtgt cctggaagct tcaggctact ccacagaggt ggtggccctg 480
297 aacaggtgac aggggtctct gcaggacatg ctgtggcagc tggacctcag cctgggtgac 540
298 tgaggccttg aaggtcactc ttctgccaag gactacgtta agggaaggaa ctctggcttc 600
299 caggtatctc caggattgaa gagcattgca ttggacacccc ttatccagga ctctgtcaat 660
300 ttccctgact cctctaagcc actcttccaa aggcataaga ccttaagcct ccttttgctt 720
301 gaaaccaaac atatatacac aggtacctat tctcaccagg aagggggtcc acccagcaaa 780

```

VERIFICATION SUMMARY

DATE: 06/25/2003

PATENT APPLICATION: US/09/804,409B

TIME: 16:19:08

Input Set : A:\0278721.APP.txt

Output Set: N:\CRF4\06252003\I804409B.raw